V-like Domain Binding Molecules

Field of the Invention

The present invention relates to V-like Domain binding molecules with affinities for target molecules. The present invention also relates to compositions comprising these V-like domain binding molecules and to methods of diagnosis or treatment which involve the use of these molecules. The present invention also relates to a method for selecting V-like Domain binding molecules with novel binding affinities and/or specificities.

Background of the Invention

Immunoglobulin Superfamily - Antibody Variable (V) Domains

Antibodies are the paradigm of specific high-affinity binding reagents and provide an antigen binding site by interaction of variable heavy (V_H) and variable light (V_L) immunoglobulin domains. The binding interface is formed by six surface polypeptide loops, termed complementarity determining regions (CDRs), three from each variable domain, which are highly variable and combined provide a sufficiently large surface area for interaction with antigen. Specific binding reagents can be formed by association of only the V_H and V_L domains into an Fv module. Bacterial expression is enhanced by joining the V-domains with a linker polypeptide into a single-chain scFv molecule. "Humanisation" of recombinant antibodies by grafting murine CDR loop structures onto a human Fv framework is disclosed by Winter et al EP-239400.

Methods to improve the expression and folding characteristics of single-chain Fv molecules were described by Nieba et al (1997). The properties of single V-domains, derived from natural mammalian antibodies, have been described by Gussow et al in WO/90/05144 and EP 0368684B1 and by Davis et al in WO/91/08482. Single camelid V-domains have been described by Hamers et al in WO/96/34103 and in WO/94/25591. A method for reducing the hydrophobicity of the surface of a human V_H domain by replacing human amino acid sequences with camelid amino acid sequences was described by Davies and Riechmann (1994). Methods to exchange other regions of human V_H sequences with camel sequences to further enhance protein stability, including the insertion of cysteine residues in CDR loops, were described by Davies and Riechmann (1996).

Several attempts to engineer high-affinity single domain binding reagents using either the V_H or V_L domains alone have been unsuccessful, due to lack of binding specificity and the inherent insolubility of single domains in the absence of the hydrophobic face where the V_H and V_L domains interact (Kortt et al, 1995).

T-cell Receptor Variable (V) Domains

The T-cell receptor has two V-domains that combine into a structure similar to the Fv module of an antibody that results from combination of the VH and VL domains. Novotny et al (1991) described how the two V-domains of the T-cell receptor (termed alpha and beta) can be fused and expressed as a single chain polypeptide and, further, how to alter surface residues to reduce the hydrophobicity directly analogous to an antibody scFv. Other publications describe the expression characteristics of single-chain T-cell receptors comprising two V-alpha and V-beta domains (Wulfing and Pluckthun,1994; Ward, 1991).

Non-antibody ligands - CTLA-4 and CD28 V-like Domains

There are a class of non-antibody ligands which bind to specific binding partners which also comprise V-like domains. These V-like domains are distinguished from those of antibodies or T-cell receptors because they have no propensity to join together into Fv-type molecules. These non-antibody ligands provide an alternative framework for the development of novel binding moieties with high affinities for target molecules. Single domain V-like binding molecules derived from these non-antibody ligands which are soluble are therefore desirable. Examples of suitable non-antibody ligands are CTLA-4, CD28 and ICOS (Hutloff et al, 1999).

Cytotoxic T-lymphocyte associated antigen 4 (CTLA-4) and the homologous cell-surface proteins CD28 and ICOS, are involved in T-cell regulation during the immune response. CTLA-4 is a 44 kDa homodimer expressed primarily and transiently on the surface of activated T-cells, where it interacts with CD80 and CD86 surface antigens on antigen presenting cells to effect regulation of the immune response (Waterhouse et al.1996, van der Merwe et al. 1997). CD28 is a 44kDa homodimer expressed predominantly on T-cells and, like CTLA-4, interacts with CD80 and CD86 surface antigens on antigen presenting cells to effect regulation of the immune response (Linsley et al.1990). Current theory suggests that competition between CTLA-4 and CD28 for available ligands controls the level of immune response, for

example, gene deletion of CTLA-4 in knock-out mice results in a massive over-proliferation of activated T-cells (Waterhouse et al. 1995).

Each CTLA-4 monomeric subunit consists of an N-terminal extracellular domain, transmembrane region and C-terminal intracellular domain. The extracellular domain comprises an N-terminal V-like domain (VLD; of approximately 14 kDa predicted molecular weight by homology to the immunoglobulin superfamily) and a stalk of about 10 residues connecting the VLD to the transmembrane region. The VLD comprises surface loops corresponding to CDR-1, CDR-2 and CDR-3 of an antibody V-domain (Metzler 1997). Recent structural and mutational studies on CTLA-4 suggest that binding to CD80 and CD86 occurs via the VLD surface formed from A'GFCC' (SEQ ID NO: 15) V-like beta-strands and also from the highly conserved MYPPPY (SEQ ID NO: 1) sequence in the CDR3-like surface loop (Peach et al. 1994; Morton et al. 1996; Metzler et al. 1997). Dimerisation between CTLA-4 monomers occurs through a disulphide bond between cysteine residues (Cys^{120}) in the two stalks, which results in tethering of the two extracellular domains, but without any apparent direct association between V-like domains (Metzler et al. 1997). Dimerisation appears to contribute exclusively to increased avidity for the ligands.

In vitro Expression of Soluble Forms of CTLA-4.

Neither the extracellular domains nor V-like domains (VLDs) of human CTLA-4 molecule have been successfully expressed as soluble monomers in bacterial cells, presumably due to aggregation of the expressed proteins (Linsley et al, 1995). Expression of the extracellular N-terminal domain (Met^{1} to Asp^{124}, comprising Cys^{120}) in *E.coli* results in production of a dimeric 28 kDa MW protein, in which two CTLA-4 V-like domains are joined by a disulphide linkage at Cys^{120}. Truncation at Val^{114} removes these cysteines and was intended to enable expression of a 14 kDa VLD in soluble, monomeric form. However, the product aggregated and it was concluded that hydrophobic sites, which were normally masked by glycosylation, were now exposed and caused aggregation (Linsley et al, 1995).

There have been some reports of successful expression of monomeric, glycosylated CTLA-4 extracellular domains in eukaryotic expression systems (ie CHO cells and the yeast Pichia pastoris; Linsley et al. 1995; Metzler et al. 1997; Gerstmayer et al. 1997). Glycosylation in these eukaryotic expression

systems is presumed to occur at the two N-linked glycosylation sites in the VLD (Asn76 and Asn108). However, high yields have only been described for expression of a gene encoding a CTLA-4 VLD fused to Ig-CH2/CH3 domains which produces a dimeric recombinant protein with 2 CTLA-4 VLDs attached to an Fc subunit (WO 95/01994 and AU 16458/95). AU 60590/96 describes mutated forms of CTLA-4 VLDs with single amino acid replacements of the first tyrosine (Y) in the MYPPPY (SEQ ID NO:1) surface loop which retain and modifies the affinity for the natural CD80 and CD86 ligands. AU 60590/96 describes the preferred soluble form of CTLA-4 VLDs as a recombinant CTLA-4/Ig fusion protein expressed in eukaryotic cells and does not solve the aggregation problem in prokaryote expression systems. EP 0757099A2 describes the use of CTLA-4 mutant molecules, for example the effect of changes on ligand binding of mutations in the CDR3-like loop.

Summary of the Invention

The present inventors have now developed novel binding molecules derived from the V-like domains (VLDs) of non-antibody ligands such as CTLA-4, CD28 and ICOS. Replacement of CDR loop structures within the VLDs results unexpectedly in the production of monomeric, correctly folded molecules with altered binding specificities and improved solubility.

Accordingly, in a first aspect the present invention provides a binding moiety comprising at least one monomeric V-like domain (VLD) derived from a non-antibody ligand, the at least one monomeric V-like domain being characterised in that at least one CDR loop structure or part thereof is modified or replaced such that the solubility of the modified VLD is improved when compared with the unmodified VLD.

Within the context of the present invention, the modification or replacement may involve any change to one or more physical characteristics (such as size, shape, charge, hydrophobicity etc) of the at least one CDR loop structure. The modification or replacement may result in a reduction in the size of the at least one CDR loop structure. In a preferred embodiment, however, at least one CDR loop structure or part thereof is modified or replaced such that

(i) the size of the CDR loop structure is increased when compared with corresponding CDR loop structure in the unmodified VLD; and/or

(ii) the modification or replacement results in the formation of a disulphide bond within or between one or more of the CDR loop structures.

In a second aspect, the present invention provides a binding moiety comprising at least one monomeric V-like domain (VLD) derived from a non-antibody ligand, the at least one monomeric V-like domain being characterised in that at least one CDR loop structure or part thereof is modified or replaced such that

(i) the size of the CDR loop structure is altered when compared with corresponding CDR loop structure in the unmodified VLD; and/or

(ii) the modification or replacement results in the formation of a disulphide bond within or between one or more of the CDR loop structures.

In a preferred embodiment of the second aspect, the size of the CDR loop structure is increased by at least two, more preferably at least three, more preferably at least six and more preferably at least nine amino acid residues.

In a further preferred embodiment, the modified binding moiety of the first or second aspect of the present invention also exhibits an altered binding affinity or specificity when compared with the unmodified binding moiety. Preferably, the effect of replacing or modifying the CDR loop structure is to reduce or abolish the affinity of the VLD to one or more natural ligands of the unmodified VLD. Preferably, the effect of replacing or modifying the CDR loop structure is also to change the binding specificity of the VLD . Thus it is preferred that the modified VLD binds to a specific binding partner which is different to that of the unmodified VLD.

The phrase "V-like domain" or "VLD" is intended to refer to a domain which has similar structural features to the variable heavy (V_H) or variable light (V_L) antibody. These similar structural features include CDR loop structures. By "CDR loop structures" we mean surface polypeptide loop structures or regions like the complementarity determining regions in antibody V-domains.

The phrase "non-antibody ligand" is intended to refer to any ligand which binds to a specific binding partner and which is not an antibody or a T-cell receptor. Examples of suitable non-antibody ligands are T-cell surface proteins such as CTLA-4, CD28 and ICOS. It will be appreciated by those skilled in the art that other non-antibody ligands which may provide V-like domains suitable for the invention are other T-cell surface proteins such as

CD2. CD4, CD7 and CD16; B cell surface proteins such as CD19, CD79a, CD22, CD33, CD80 and CD86; adhesion molecules such as CD48, CD541CAM and CD58. These molecules, which are listed in Table 1, provide a non-exhaustive list of structures which may form the basis for the single domain binding molecules of the present invention.

The phrase "V-like domain derived from a non-antibody ligand" is intended to encompass chimeric V-like domains which comprise at least part of a V-like domain derived from a non-antibody ligand.

TABLE 1: NON-ANTIBODY LIGANDS

Molecule	Size	Structure
Tcell Surface Proteins		
CD2	45-58kDa	VC[1] domains
CD4	55kDa	V2C2
CD7	40kDa	V domain
CD16	50-65kDa	2x C domains
B cell Surface Proteins		
CD19	95kDa	2x C domains
CD79a	33kDa	
CD22	130-140kDa	1xV 6xC domains
CD33	67kDa	VC domain
CD80	60kDa	VC domain
CD86	60kDa	VC domain
Adhesion molecules		
CD48	45kDa	VC domain
CD541CAM	85-110kDa	
CD58	55-70kDa	VC domain

1 V = variable Ig domain, C = constant domain

These molecules are discussed in (1) The Leucocyte Antigen Facts Book, 1993, Eds Barclay et al., Academic Press, London; and (2) CD Antigens 1996 (1997) Immunology Today 18, 100-101, the entire contents of which are incorporated herein by reference.

The "solubility" of modified binding moieties of the present invention correlates with the production of correctly folded, monomeric domains. The solubility of the modified VLDs may therefore be assesed by HPLC. For example, soluble (monomeric) VLDs will give rise to a single peak on the HPLC chromatograph, whereas insoluble (eg. multimeric and aggregated) VLDs will give rise to a plurality of peaks. A person skilled in the art will therefore be able to detect an increase in solubility of modified VLDs using routine HPLC techniques.

It will be appreciated that the binding moieties of the present invention may be coupled together, either chemically or genetically, to form multivalent or multifunctional reagents. For example, the addition of C-terminal tails, such as in the native CTLA-4 with Cys^{120}, will result in a dimer.

The binding moieties of the present invention may also be coupled to other molecules for various diagnostic formulations. For example, the VLDs may comprise a C-terminal polypeptide tail or may be coupled to streptavidin or biotin for multi-site *in vitro* assays. The VLDs may also be coupled to radioisotopes, dye markers or other imaging reagents for *in vivo* detection and/or localisation of cancers, blood clots, etc. The VLDs may also be immobilised by coupling onto insoluble devices and platforms for diagnostic and biosensor applications.

In a most preferred embodiment of the first aspect of the present invention, the V-like domain is derived from the extracellular domain of the CTLA-4 molecule or the CD28 molecule. In a further preferred embodiment one or more surface loops of the CTLA-4 V-like domain and preferably the CDR-1, CDR-2 or CDR-3 loop structures are replaced with a polypeptide which has a binding affinity for a target molecule of interest. Target molecules of interest comprise, but are not limited to, drugs, steroids, pesticides, antigens, growth factors, tumour markers, cell surface proteins or viral coat proteins. It will be appreciated that these VLDs may be polyspecific, having affinities directed by both their natural surfaces and modified polypeptide loops.

In a further preferred embodiment the effect of replacing or modifying the CTLA-4, CD28 and ICOS V-like domain surface loops is to abolish the natural affinity to CD80 and CD86.

In one preferred embodiment, one or more of the CDR loop structures of the VLD are replaced with one or more CDR loop structures derived from an antibody. The antibody may be derived from any species. In a preferred embodiment, the antibody is derived from a human, rat, mouse, camel, llama or shark. The antibody or antibodies may be selected from the camel antibody cAB-Lys3 and the human anti-melanoma antibody V86.

In a further preferred embodiment, one or more of the CDR loop structures are replaced with a binding determinant derived from a non-antibody polypeptide. For example, one or more of the CDR loop structures may be replaced with a polypeptide hormone, such as somatostatin which is a 14 residue intra-disulphide bonded polypeptide important in cancer cell recognition, or with a viral protein such as the human influenza virus haemagglutinin protein.

In a further preferred embodiment the V-like domain of the binding moiety comprises CDR loop structures homologous in character to CDR loop structures found in camelid or llama antibodies. For example, the CDR loop structures may contain one or more non-conventional substitutions (eg. hydrophobic to polar in nature). In another preferred embodiment, the CDR-1 and CDR-3 loop structures may adopt non-canonical conformations which are extremely heterologous in length. The V-like domain may also possess a disulphide linkage interconnecting the CDR-1 and CDR-3 loop structures (as found in some camel V_HH antibodies) or the CDR-2 and CDR-3 loop structures (as found in some llama V_HH antibodies).

In a third aspect the present invention provides a polynucleotide encoding a binding moiety of the first or second aspect of the present invention. The polynucleotide may be incorporated into a plasmid or expression vector.

In a fourth aspect the present invention provides a prokaryotic or eukaryotic host cell transformed with a polynucleotide according to the third aspect of the present invention.

In a fifth aspect the present invention provides a method of producing a binding moiety which comprises culturing a host cell according to the fourth aspect of the present invention under conditions enabling expression of the binding moiety and optionally recovering the binding moiety.

In a preferred embodiment of the present invention the binding moiety is produced by expression in a bacterial host. Preferably, the binding moiety is unglycosylated.

In a sixth aspect the present invention provides a pharmaceutical composition comprising a binding moiety of the first or second aspect of the present invention and a pharmaceutically acceptable carrier or diluent.

In a seventh aspect the present invention provides a method of treating a pathological condition in a subject, which method comprises administering to the subject a binding moiety according to the first or second aspect of the present invention.

For *in vivo* applications it is preferable that VLDs are homologous to the subject of treatment or diagnosis and that any possible xenoantigens are removed. Accordingly it is preferred that VLD molecules for use in clinical applications are substantially homologous to naturally occurring human immunoglobulin superfamily members.

In an eighth aspect the present invention provides a method of selecting a binding moiety with an affinity for a target molecule which comprises screening a library of polynucleotides for expression of a binding moiety with an affinity for the target molecule, the polynucleotides encoding VLDs derived from one or more non-antibody ligands, wherein the polynucleotides have been subjected to mutagenesis which results in a modification or replacement in at least one CDR loop structure in at least one VLD and wherein the solubility of the isolated modified VLD is improved when compared with the isolated unmodified VLD.

It will be appreciated by those skilled in the art that within the context of the eighth aspect of the present invention, any method of random or targetted mutagenesis may be used to introduce modifications into the V-like domains. In a preferred embodiment, the mutagenesis is targetted mutagenesis. Preferably, the targetted mutagenesis involves replacement of at least one sequence within at least one CDR loop structure using splice overlap PCR technology.

It will also be appreciated by those skilled in the art that the polynucleotide library may contain sequences which encode VLDs comprising CDR loop structures which are substantially identical to CDR loop structures found in naturally occurring immunoglobulins as well as

sequences which encode VLDs comprising non-naturally occurring CDR loop structures.

In a preferred embodiment of the eighth aspect of the present invention, the screening process involves displaying the modified V-like domains as gene III protein fusions on the surface of bacteriophage particles. The library may comprise bacteriophage vectors such as pHFA, fd-tet-dog or pFAB.5c containing the polynucleotides encoding the V-like domains.

In a further preferred embodiment of the eighth aspect, the screening process involves displaying the modified V-like domains in a ribosomal display selection system.

Throughout this specification, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

Brief Description of the Drawings

Figure 1: CTLA-4 VLD-Specific Oligonucletides. (SEQ ID NO: 2-56)

Figure 2: Polynucleotide sequence of complete cDNA encoding human CTLA-4 and polypeptide sequence of the VLD of human CTLA-4.

Figure 3: Display of CTLA-4 VLD STMs as gene 3 fusions on the surface of phage or phagemid. CTLA-4 VLD STMs are depicted as black spheroids; gene 3 protein is depicted as white spheroids; FLAG polypeptide is depicted in grey; genes are marked in a similar colour code and are depicted in an oval phage(mid) vector.

Figure 4: Schematic representation of the somatostatin polypeptide. Somatostatin (somatotropin release-inhibiting factor SRIF) is a cyclic 14-amino acid polypeptide. The cyclic nature is provided by a disulphide linkage between the cysteine residues at positions 3 and 14. The four residues which constitute the tip of the loop (Phe-Trp-Lys-Thr) are (SEQ ID NO: 140) implicated in binding to members of the somatostatin receptor family.

Figure 5: Size exclusion HPLC profiles of affinity purified CTLA-4 VLD and CTLA-4-Som3 STM. Recombinant human CTLA-4 proteins were expressed in E. coli host TG1 from vector pGC, purified from periplasmic extracts by anti-FLAG affinity chromatography and subjected to size exclusion chromatography on a calibrated Superose 12 HR column. The elution profiles of purified CTLA-4 VLD and CTLA-4-Som3 STM are overlayed in this graph. CTLA-4 VLD comprises tetramer (21.86 min), dimer (26.83) and monomer (29.35 min). CTLA-4-Som3 STM comprises dimer (26.34) and monomer (29.28). Traces represent absorbance at 214 nm and are given in arbitrary units.

Figure 6: Schematic diagram of CTLA-4 VLD loop replacements. Construct A (CTLA-4 VLD: S2) represents the wild-type CTLA-4 extracellular V-domain, spanning residues 1-115. Constructs B (CTLA-4-Som1: PP2) and C (CTLA-4-Som1-Cys120; PP5) both contain the 14 residue somatostain polypeptide in CDR1. PP5 also carries a C-terminal extension containing Cys120. Construct D (CTLA-4-Som3; PP8) contains the 14 residue somatostatin polypeptide in place of CDR3. In construct E (CTLA-4-HA2: XX4), CDR2 has been replaced with a haemagglutinin tag. In construct F (CTLA-4-Som1-Som3: VV3), both CDR1 and CDR3 have been replaced with the somatostain polypeptide. In construct G (CTLA-4-Som-HA2-Som3: ZZ3) CDR1 and CDR3 are replaced with the somatostatin polypeptide whilst CDR2 is replaced with a haemagglutinin tag. In construct H (CTLA-4-anti-lys:2V8), all three CDR loop structures have been replaced with the CDR loops from a camel anti-lysozyme V_HH molecule. Construct I (CTLA-4-anti-mel: 3E4) represents CTLA-4 VLD in which all three CDRs have been replaced by the VH CDR loops from anti-melanoma antibody V86 (Cai And Garen, 1997). PelB, cleavable pectate lyase secretion sequence (22 aa); flag, dual flag tag (AAADYKDDDDKAADYKDDDDK).

Figure 7: HPLC profiles of purified recombinant human CTLA-4 STMs. Recombinant CTLA-4 VLDs were expressed in E. coli host TG1 from vector pGC, purified from periplasmic extracts by anti-FLAG affinity chromatography and subjected to size exclusion chromatography on a calibrated Superose 12 HR column. The elution profiles of the purified proteins are shown. Panel A, CTLA-4 DIMER (PP5); Panel B, CTLA-4R (S2);

Panel C, CTLA-4-HA2 (XX4); Panel D, CTLA-4-Som3 (PP8); Panel E, CTLA-4-Som1 (PP2); Panel F, CTLA-4-Som1-Som3 (VV3); Panel G, CTLA-4-Som-HA2-Som3 (ZZ3); Panel H, CTLA-4-anti-lys (2V8); Panel I, CTLA-4-anti-mel (3E4).). Traces represent absorbance at 214 nm and are given in arbitrary units.

Figure 8: Comparison by size exclusion FPLC analysis of affinity purified CTLA-4 constructs synthesised using bacterial expression vector pGC or pPOW. Recombinant human CTLA-4R or its loop variants were expressed in E. coli host TOP10F', purified from periplasmic extracts by anti-FLAG affinity chromatography and subjected to size exclusion chromatography on a calibrated Superose 12 HR column. The elution profiles of proteins expressed from vector pGC are shown on the left, whilst proteins expressed from vector pPOW are shown on the right. Panel A, wild-type CTLA-4 VLD (S2); B, CTLA-4-Som1(PP2); C, CTLA-4-Som3(PP8); D, CTLA-4-Anti-lys(2V8); E, CTLA-4-Som1-HA2-Som3(ZZ3).

Figure 9: Protein stability analysis. Stability of monomer preparations of CTLA-4 VLD and loop variant constructs was analysed by size exclusion fplc chromatography on a precalibrated superose 12 hr (Pharmacia) column following several cycles of freeze/thawing. Aliquots of each protein were tested immediately after peak purification and following two cycles of freeze/thawing. A, CTLA-4 VLD (S2); B, CTLA-4-Som1 (PP2); C, CTLA-4-Som3 (PP8); D, CTLA-4-anti-lys (2V8); E, CTLA-4-Som-HA2-Som3 (ZZ3).

Figure 10: Lysozyme binding characteristics of CTLA-4-anti-lys construct 2V8.

Figure 11: Screening of CTLA-4 VLD phagemid library on immobilised Sh bleomycin.

Figure 12: Screening of CTLA-4 VLD libraries in solution

Detailed Description of the Invention

The present invention relates to the design of novel soluble VLD binding molecules derived from the V-like domain of immunoglobulin

superfamily members, such as the human CTLA-4 molecule. The preferred binding molecules of the present invention provide the following advantages (i) use of a native human protein obviates the need for subsequent humanisation of the recombinant molecule, a step often required to protect against immune system response if used in human treatment; (ii) the domain is naturally monomeric as described above (incorporation of residue Cys120 in a C-terminal tail results in production of a dimeric molecule); and (iii) structural modifications have resulted in improved E.coli expression levels.

Prior to publication of the first CTLA-4 structure determination, available sequence data and mutational analyses of both this molecule and CD28 were analysed. This allowed modelling and prediction of the regions corresponding to antibody CDR1, 2 and 3 regions. It was hypothesised that such areas would be susceptible to mutation or substitution without substantial effect upon the molecular framework and hence would allow expression of a correctly folded molecule. The subsequently published structure (Metzler et al. 1997) showed these predictions to be accurate, despite the unexpected separation of CDR1 from the ligand-binding site, and the extensive bending of CDR3 to form a planar surface contiguous with the ligand binding face.

In an initial set of experiments the V-like domain of the human CTLA-4 molecule was modified by replacement of CDR loop structures with either of two defined polypeptides. The two polypeptides were human somatostatin (Som) and a portion of the human influenza virus haemagglutinin protein (HA-tag). Somatostatin (SRIF: somatotropin release-inhibiting factor) is a 14 residue polypeptide comprising a disulphide bond that forces the central 10 residues into a loop. Human somatostatin is biologically widespread within the body and mediates a number of diverse physiological functions such as regulation of growth hormone secretion etc (Reisne, 1995). Human somatostatin binds a number of specific receptors (there are at least five subtypes) which have differing tissue specificities and affinities (Schonbrunn et al. 1995). These aspects of binding and activation are as yet poorly understood, but one salient feature is the high density of somatostatin receptors present on a number of cancerous cell lines, for example cancers of the neuro-endocrine system and small lung cancers (Reubi 1997). Artificial analogues of somatostatin have been produced for

imaging of such tumours which are resistant to degradation compared with the highly labile somatostatin polypeptide.

The haemagglutinin epitope sequence consists of the 9 residues YPYDVPDYA. A commercially produced antibody is available which specifically recognises this sequence. The epitope tag can be detected when randomly or directionally incorporated within the structure of proteins (Canfield et al. 1996).

Replacement of one or more CDR loop structures in the CTLA-4 V-like domain with somatostatin or the HA-tag resulted in the production of soluble, monomeric, unglycosylated binding molecules using different bacterial expression systems. This surprising finding shows that V-like domains provide a basic framework for constructing soluble, single domain molecules, where the binding specificity of the molecule may be engineered by modification of the CDR loop structures.

The basic framework residues of the V-like domain may be modified in accordance with structural features present in camelid antibodies. The camel heavy chain immunoglobulins differ from "conventional" antibody structures by consisting of only a single VH domain (Hamers-Casterman et al. 1993). Several unique features allow these antibodies to overcome the dual problems of solubility and inability to present a sufficiently large antigen binding surface.

First, several non-conventional substitutions (predominantly hydrophobic to polar in nature) at exposed framework residues reduce the hydrophobic surface, while maintaining the internal beta-sheet framework structure (Desmyter et al. 1996). Further, within the three CDR loops several structural features compensate for the loss of antigen binding-surface usually provided by the VL domain. While the CDR2 loop does not differ extensively from other VH domains, the CDR-1 and -3 loops adopt non-canonical conformations which are extremely heterologous in length. For example, the H1 loop may contain anywhere between 2-8 residues compared to the usual five in Ig molecules. However, it is the CDR3 loop which exhibits greatest variation: in 17 camel antibody sequences reported, the length of this region varies between 7 and 21 residues (Muyldermans et al. 1994). Thirdly, many camelid VH domains possess a disulphide linkage interconnecting CDRs -1 and -3 in the case of camels and interconnecting CDRs -1 and -2 in the case of llamas (Vu et al. 1997). The function of this structural feature appears to

be maintenance of loop stability and providing a more contoured, as distinct from planar, loop conformation which both allows binding to pockets within the antigen and gives an increased surface area. However, not all camelid antibodies possess this disulphide bond suggesting that it is not an absolute structural requirement.

These foregoing features have enabled camelid V-domains to present as soluble molecules in vivo and with sufficiently high affinity to form an effective immune response against a wide variety of target antigens. In contrast, cell surface receptors of the Ig superfamily (such as CD4 and CD2) comprise V-like binding domains and appear to bind cognate receptors with surface features other than the CDR loops. These V-like domains bind to cognate receptors with very low affinity. Physiological signalling between two cells are mediated by the avidity of multi-point binding, when two cell surfaces connect and each contains multiple receptors. CD2 is a well-characterised example: binding to CD58 is mediated by a highly constrained set of minor electrostatic interactions generated by charged and polar residues located in the GFCC'C" face (SEQ ID NO:141) (not the antibody type CDR-1, CDR-2 or CDR-3 loops). This results in a low affinity but highly specific interaction (Bodian et al 1994).

The present invention also relates to a method for generating and selecting single VLD molecules with novel binding affinities for target molecules. This method involves the application of well known molecular evolution techniques to V-like domains derived from members of the immunoglobulin superfamily. The method may involve the production of phage or ribosomal display libraries for screening large numbers of mutated V-like domains.

Filamentous fd-bacteriophage genomes are engineered such that the phage display, on their surface, proteins such as the Ig-like proteins (scFv, Fabs) which are encoded by the DNA that is contained within the phage (Smith, 1985; Huse et al., 1989; McCafferty et al., 1990; Hoogenboom et al., 1991). Protein molecules can be displayed on the surface of Fd bacteriophage, covalently coupled to phage coat proteins encoded by gene III, or less commonly gene VIII. Insertion of antibody genes into the gene III coat protein gives expression of 3-5 recombinant protein molecules per phage, situated at the ends. In contrast, insertion of antibody genes into gene VIII has the potential to display about 2000 copies of the recombinant protein per

phage particle, however this is a multivalent system which could mask the affinity of a single displayed protein. Fd phagemid vectors are also used, since they can be easily switched from the display of functional Ig-like fragments on the surface of Fd-bacteriophage to secreting soluble Ig-like fragments in E. coli. Phage-displayed recombinant protein fusions with the N-terminus of the gene III coat protein are made possible by an amber codon strategically positioned between the two protein genes. In amber suppressor strains of E. coli, the resulting Ig domain-gene III fusions become anchored in the phage coat.

A selection process based on protein affinity can be applied to any high-affinity binding reagents such as antibodies, antigens, receptors and ligands (see, for example, Winter and Milstein, 1991, the entire contents of which are incorporated herein by reference). Thus the selection of the highest affinity binding protein displayed on bacteriophage is coupled to the recovery of the gene encoding that protein. Ig-displaying phage can be affinity selected by binding to cognate binding partners covalently coupled to beads or adsorbed to plastic surfaces in a manner similar to ELISA or solid phase radioimmunoassays. While almost any plastic surface will adsorb protein antigens, some commercial products are especially formulated for this purpose, such as Nunc Immunotubes.

Ribosomal display libraries involve polypeptides synthesised de novo in cell-free translation systems and displayed on the surface of ribosomes for selection purposes (Hanes and Pluckthun, 1997; He and Taussig, 1997). The "cell-free translation system" comprises ribosomes, soluble enzymes required for protein synthesis (usually from the same cell as the ribosomes), transfer RNAs, adenosine triphosphate, guanosine triphosphate, a ribonucleoside triphosphate regenerating system (such as phosphoenol pyruvate and pyruvate kinase), and the salts and buffer required to synthesize a protein encoded by an exogenous mRNA. The translation of polypeptides can be made to occur under conditions which maintain intact polysomes, i.e. where ribosomes, mRNA molecule and translated polypeptides are associated in a single complex. This effectively leads to "ribosome display" of the translated polypeptide.

For selection, the translated polypeptides, in association with the corresponding ribosome complex, are mixed with a target molecule which is bound to a matrix (e.g. Dynabeads). The target molecule may be any

compound of interest (or a portion thereof) such as a DNA molecule, a protein, a receptor, a cell surface molecule, a metabolite, an antibody, a hormone or a virus. The ribosomes displaying the translated polypeptides will bind the target molecule and these complexes can be selected and the mRNA re-amplified using RT-PCR.

Although there are several alternative approaches to modify binding molecules the general approach for all displayed proteins conforms to a pattern in which individual binding reagents are selected from display libraries by affinity to their cognate receptor. The genes encoding these reagents are modified by any one or combination of a number of *in vivo* and *in vitro* mutation strategies and constructed as a new gene pool for display and selection of the highest affinity binding molecules.

In order that the nature of the present invention may be more clearly understood preferred forms thereof will now be described with reference to the following examples.

Example 1

Gene Construction and Cloning

CTLA-4 STM (STM: soluble truncated mutants of CTLA-4, used herein to describe CTLA-4 chimaeric V-like domain proteins) gene construction and cloning was by standard and well-described techniques (Polymerase chain reaction with specifically designed oligonucleotide primers, splice overlap extension, restriction enzyme digests etc). A list of oligonucleotide primers used is given in Figure 1.

The wild-type STM construct was amplified from cloned human CTLA-4 DNA (Figure 2) (and could be similarly amplified from reverse transcribed human cDNA by a competent worker in the field) using the oligonucleotide primers #3553 and #4316, which incorporated SfiI and NotI restriction sites at the 5' and 3' ends respectively. These terminal primers were used in all further constructions except: (i) where #4851 or #5443 was used to incorporate an ApaL1 site at the 5' end; (ii) where #4486 was used to add a C-terminal tail including residue Cys120; (iii) where #5467 was used to incorporate an EcoR1 site at the 5' end; and (iv) where the specific set of extension primers were used for ribosomal display.

A splice overlap PCR strategy using combinations of the oligonucleotides primers listed in Figure 1 was used to produce variations of

CDR-1, CDR-2 and/or CDR-3 loop structure replacements. The variations, which are described in greater detail in the following examples are listed in Table 2.

TABLE 2

CDR-1 combinations

	CDR-1	
CTLA-4 VLD	S^{19}FVCEYA.**SPGKATE**.....	VRV...
Anti-lysozyme	S^{19}FVCEYA.**SGYTIGPYCMG**.....	VRV...
Somatostatin-14	S^{19}FVCEYA.**AGCKNFFWKTFTSCATE**.	VRV...
Anti-melanoma	S^{19}FVCEYA.**SGFTFSSYAMS**.....	VRV...
~~Randomisation 1~~	S^{19}FVCEYA.**XXXXXXXG**.....	VRV...
~~Randomisation 2~~	S^{19}FVCEYA.**XXXXXXXXXCXG**.....	VRV...
~~Randomisation 3~~	S^{19}FVCEYA.**XXarXarXXarCXG**.....	VRV...
~~Randomisation 4~~	S^{19}FVCEYA.**SPGXXXX**.....	VRV...
~~Randomisation 5~~	S^{19}FVCEYA.**SPGXCXX**.....	VRV...
~~Randomisation 6~~	S^{19}FVCEYA.**XXXXXXXXXATE**.....	VRV...
~~Randomisation 7~~	S^{19}FVCEYA.**XXXXXXCXATE**.....	VRV...
~~Randomisation 8~~	S^{19}FVCEYA.**AGCKNXXXXXXTSCATE**.	VRV...

CDR-2 combinations

	CDR-2	
CTLA-4 VLD	Q^{44}VTEVCAA.**TYMMGNELTF**.LDDSICT...	
Anti-lysozyme	Q^{44}VTEVCAA.**AINMGGGITF**.LDDSICT...	
Haemagglutinin tag	Q^{44}VTEVCAA.**TYPYDVPDYA**.LDDSICT...	
Anti-melanoma	Q^{44}VTEVCAA.**AISGSGGSTY**.LDDSICT...	
~~Randomisation 1~~	Q^{44}VTEVCAA.**TYXXGXELTF**.LDDSICT...	
~~Randomisation 2~~	Q^{44}VTEVCAA.**CYXXGXELTF**.LDDSICT...	

CDR-3 combinations

	CDR-3	
CTLA-4 VLD (SEQ ID no:89)	C^{93}KV.ELMYPPPYYL.....	GIG...
Anti-lysozyme (SEQ ID no:90)	C^{93}KV.DSTIYASYYECGHGLSTGGYGYDS.	GIG...
Somatostatin-14 (SEQ ID no:91)	C^{93}KV.EAGCKNFFWKTFTSC.....	GIG...
Anti-melanoma (SEQ ID no:92)	C^{93}KV.GWGLRGEEGDYYMDV.....	GIG...
~~Randomisation 1~~ Randomization 1 (SEQ ID no:93)	C^{93}KV.XXXXXXXXXXX.....	GIG...
~~Randomisation 2~~ Randomization 2 (SEQ ID no:94)	C^{93}KV.XXXXXXXXXXXXXXX.....	GIG...
~~Randomisation 3~~ Randomization 3 (SEQ ID no:95)	C^{93}KV.XXXXXX.C.XXXX.....	GIG...
~~Randomisation 4~~ Randomization 4 (SEQ ID no:96)	C^{93}KV.XXXXXXX.C.XXXX.....	GIG...
~~Randomisation 5~~ Randomization 5 (SEQ ID no:97)	C^{93}KV.XXXXXXXX.C.XXXXX.....	GIG...
~~Randomisation 6~~ Randomization 6 (SEQ ID no:98)	C^{93}KV.XXXXXXXXXX.C.XXXXXXX.....	GIG...
~~Randomisation 7~~ Randomization 7 (SEQ ID no:99)	C^{93}KV.EXXXXXXXXX.....	GIG...
~~Randomisation 8~~ Randomization 8 (SEQ ID no:100)	C^{93}KV.EXXXXXX.CXXXXXXX.....	GIG...
~~Randomisation 9~~ Randomization 9 (SEQ ID no:101)	C^{93}KV.EAGCKNXXXXXXTSC.....	GIG...

For generation of randomised sections of the CDR loop structures, similar splice-overlap techniques were used with oligonucleotides where a given triplet(s) were encoded by the sequence NNg/T where N represents any of the four possible nucleotide bases. This combination covers all possible amino acid residues. Alternatively, randomisation was biased towards certain subsets of amino acids (for example aromatic residues, Figure 1, #5452).

In some instances, a variant technique was used for STM gene construction, where randomised oligonucleotide primers were designed which incorporated restriction sites for direct cloning into the similarly modified (with complementary restriction sites) CTLA-4 VLD framework (for example Figure 1, #4254).

Completed constructs were cut with approriate combinations of restriction enzymes (for example Sfi1,Not1, ApaL1, EcoR1) and cloned into like sites in appropriate expression vectors. These vectors comprise: (i) for production of soluble protein expression vectors pGC (Coia et al, 1996) and pPOW (Power et al, 1992; Kortt et al. 1995) (ii) for bacteriophage and

phagemid display, completed STM constructs were cut with the restriction enzymes SfiI and NotI or ApaLI and NotI and cloned into the vectors pHFA, and pFAB.5c (phagemid) or pfd-Tet-DOG (phage). These vectors allow display of the STMs as gene3 protein fusions on the surface of bacteriophage in 1-2 (phagemid) or 3-5 (phage) copies per bacteriophage particle (Figure 3).

All DNA constructs were verified by restriction analysis and DNA sequencing and tested for expression of recombinant protein by standard and well-understood techniques (Polyacrylamide gel electrophoresis, Western blot etc).

Example 2

Production and Isolation of Recombinant STM Proteins

Recombinant proteins were produced using vectors which represent different protocols for periplasmic expression systems. These vectors were (i) pGC: this vector allows high level expression of heterologous proteins by chemical (IPTG) induction, which are targeted to the periplasmic space by means of a leader sequence. The leader sequence is subsequently cleaved to produce the mature protein. In addition, this vector contains two in-frame 8 residue tag sequences (FLAG tags) which allow affinity purification of the recombinant protein. (ii) pPOW, which, like pGC, allows high level heat inducible expression of proteins targeted to the periplasmic space by means of a cleavable leader sequence and two in-frame 8 residue tag sequences (FLAG tags).

Recombinant proteins were purified by the following methods. which are but two variations of well established techniques. (i) Bacterial clones in vector pGC were grown overnight in 2YT medium/37°C /200 rpm/100mg/ml ampicillin, 1% glucose (final). Bacteria were diluted 1/100 into either 0.5 or 2l of 2YT medium supplemented with 100mg/ml ampicillin, 0.1% glucose (final), and grown at 28°C/ 200 rpm. These cultures were grown to an optical density of between 0.2-0.4, at which stage they were induced with 1mM IPTG (final). Cultures were allowed to grow for 16 hours (overnight) before harvesting. Bacteria were collected by centrifugation (Beckman JA-14 rotor or equivalent/6K/10min/4°C) and the periplasmic fraction collected by standard techniques. Briefly, this involved resuspension of bacterial pellets in a 1/25th volume of spheroplast forming buffer consisting of 100mM Tris-HCl/0.5M sucrose/0.5 mM EDTA (pH8.0), followed by addition of 1/500th

volume of hen egg lysozyme (2mg/ml in water) and incubation for 10min. A 0.5x solution of the above spheroplasting buffer was then added to a final volume of 1/5th of the original culture, and the incubation continued for a further 20min. The cell debris was then pelleted by centrifugation (Beckman JA-14 rotor or equivalent/9K/15min/4°C) and the supernatant containing the periplasmic fraction collected. All of the above procedures were performed at 4°C. Samples were processed immediately by sonication, filtration through a 0.45μ nitrocellulose membrane and processed immediately or stored at 4°C in the presence of sodium azide (0.05%). If freezing was required, no more than one freeze-thaw cycle was allowed. (ii) Bacterial clones in pPOW were grown overnight at 30°C in 100 ml 2xYT broth containing 100 μg/ml (w/v) ampicillin. On the following day cultures were used to inoculate 900 ml fresh 2xYT broth containing 100 μg/ml (w/v) ampicillin, to OD600nm = 0.2-0.5, and grown at 30°C with shaking (150-200 rpm) until OD600nm = 4 i.e. late log phase. To induce recombinant protein expression, the temperature was raised to 42°C for 1 hour and then dropped to 20°C for a further hour. Cells were harvested by centrifugation (Beckman JA-14 /6K rpm/5 min/4°C), the cell pellet resuspended in 100 ml extraction buffer (20mM Tris pH 8.0/ 0.2mg/ml (w/v) lysozyme/0.1% (v/v) Tween-20) and incubated at 4°C overnight. Samples were sonicated for 30 seconds and cellular debris collected by centrifugation (Beckman JA-14 /14K rpm/10 min/4°C). The aqueous phase, containing the "lysozyme" wash, was retained. Cell pellets were then washed twice with ice-cold water and this "osmotic shock" wash was retained. Each wash consisted of resuspending the pellet in 100 ml ice-cold water followed by incubation on ice for 10 minutes in the first instance followed by 1 hour in the second instance. Following centrifugation (Beckman JA-14 /14K rpm/10 min/4°C), the pH of the aqueous phase was adjusted by addition of 10 ml 10xTBS, pH 8. The "lysozyme" and "osmotic shock" washes were pooled and constitute the soluble or "periplasmic" protein fraction. Periplasmic fractions were sonicated, filtered through a 0.45μ nitrocellulose membrane and processed immediately or stored at 4°C in the presence of sodium azide (0.05%), PMSF (23 μg/ml) and EDTA (50 mM).

Recombinant proteins were purified by affinity chromatography through a divinyl sulphone activated agarose (Mini-Leak)-linked anti-FLAG antibody column. Periplasmic extracts were directly loaded onto a 10 ml

anti-FLAG column which had been pre-equilibrated in TBS (pH 8) containing 0.05% (w/v) sodium azide. Bound proteins were eluted with Immunopure Gentle Ag/Ab Elution Buffer (Pierce). Samples were then dialysed against TBS/0.05% (w/v) azide (pH 8), concentrated by ultrafiltration over a 3 kDa cut-off membrane (YM3, Diaflo), and analysed by HPLC on a pre-calibrated Superose 12 HR or Superdex 200 HR column (Pharmacia Biotech), at a flow rate of 0.5 ml/min. Fractions corresponding to monomeric , dimeric and tetrameric species were collected, concentrated as above, and stored at 4°C prior to analysis. Protein concentration was determined spectrophotometrically using an extinction coefficient at A280 of 1.27 for the CTLA-4R extracellular domain, 0.92 for CTLA-4-Som1, 1.13 for CTLA-4-Som3, 1.05 for CTLA-4-Anti-Lys. All of the above protein chemistry methods are standard techniques within the field. Purified proteins were analysed by standard techniques for example polyacrylamide gel electrophoresis, western blot, dot blot etc.

Cloning and expression in the bacteriophage expression vectors pHFA, pFAB.5c and fd-tet dog, and subsequent production of recombinant bacteriophage, were by standard and well-established techniques. Screening of libraries of randomised CTLA-4 STMs was by standard and well-established techniques (Galanis et al 1997).

Example 3

CTLA-4 STMs incorporating Somatostatin and Haemagglutinin Peptides.

Initially the CDR1 or the CDR3 loop structures of the CTLA-4 STM were replaced with the somatostatin polypeptide. This 14 residue polypeptide is conformationally constrained by an intra-disulphide linkage between Cys3 and Cys14 (Figure 4). This was reasoned to form a discrete protein loop, analogous to the CDR loops found in antibodies, and particularly analogous to the long CDRs found in camelid antibodies which are also stabilised by a disulphide linkage. The effect of substituting CDR1 in the presence or absence of Cys120 ie. whether a dimer could be produced, was also tested. These experiments produced an unexpected and surprising result. Substitution of either CDR-1 or -3 with somatostatin significantly enhanced the production of monomeric protein. This is illustrated in Figure 5 where replacement of the CDR-3 loop structure with somatostatin

significantly increased the ratio of monomeric to dimeric/tetrameric protein species.

In further experiments, simultaneous replacement of both CDR-1 and -3 loop structures by somatostatin resulted in production of high-levels of monomeric protein. This shows that extensive loop structure replacements can be accommodated by the CTLA-4 scaffolding. Structurally, one of the somatostatin loops may lie flat against the face of the molecule in a manner analogous to that of the CDR-3 loop structure of CTLA-4 VLD.

In a further extension of the CDR loop structure-replacement strategy, a region corresponding to CDR-2 was replaced with the 8-residue haemagglutinin (HA) tag sequence. Use of the conformationally constrained somatostatin loop in this position was considered unsuitable as this region partially encompasses the C" beta strand running the length of the molecule. The HA tag could be detected upon this CTLA-4 STM by use of an anti-HA antibody. Gel filtration experiments showed the presence of a range of protein species, from monomeric through to aggregated species suggesting that CDR-2-only substitutions were not stable (Figures 6,7).

Final proof of principle that the CTLA-4 CDR loop structures could be replaced with other polypeptides to produce monomeric, soluble, STMs was by simultaneous replacement of all three CDR loop structures with two somatostatin and one HA epitope respectively. This STM produced a correctly folded and monomeric protein upon gel filtration chromatography (Figures 6,7).

The positions of CDR loop structure substitutions within the CTLA-4 VLD for the various STMs are shown in figure 6. HPLC profiles of affinity-purified STM proteins are shown in figure 7. Identical results were obtained for proteins produced in two different protein expression systems: pGC where protein expression is chemically induced, and pPOW where protein expression is temperature induced (see Example 2)(Figure 8). Polyacrylamide gel electrophoresis followed by western blot analysis indicated that the CTLA-4 STMs could be reduced and ran at the expected molecular weights and absent of glycosylation. Testing of isolated monomeric STM proteins showed that they remained monomeric after zero, one, or two freeze-thaw cycles (figure 9).

CTLA-4 STMs retained the correct conformation since a conformationally-specific anti-CTLA-4 antibody recognised STMs with both

CDR1 and -3 loop structure replacements. Interestingly, this antibody recognised the wild type monomer and the dimer (CDR1 replaced) poorly, contrasting with the strong reaction observed for the modified protein species. This suggests that in the wild type STM some form of local interaction is occurring that occludes the antibody binding site, and that this interaction is similar to the result when two CTLA-4 molecules are tethered together (presumably blocking access to the antibody).

Example 4

CTLA-4 STMs Based Upon a Camel anti-Lysozyme Antibody.

The camel V_HH antibody cAb-Lys3 isolated from immunised camels specifically binds within the active site cleft of hen egg lysozyme (Desmyter et al. 1996). To illustrate the ability of CTLA-4 STMs to function in a similar fashion, the three CDR loop structures of CTLA-4 VLD STM were replaced with the three CDR loop regions from cAb-Lys3. Positions and sequence of the substitutions are shown in figure 6. Expression of this STM (2V8) in either pGC or pPOW based expression systems resulted in production of predominantly monomeric soluble protein (Figures 7, 8). Protein solubility of this CTLA-4 STM was superior to native CTLA-4 VLD. ELISA analysis showed that (pGC produced) purified monomeric protein specifically bound hen egg lysozyme compared to non-specific antigens and compared to the CTLA-4 STM with somatostatin substituted within the CDR1 loop structure (PP2) (Figure 10A). Real-time binding analysis by BIAcore showed that the lysozyme specifically bound to immobilised anti-lysozyme STM (Figure 10B). The CTLA-4 STM framework is thus folding correctly and presenting the CDR loop structures in a manner in which they can bind lysozyme antigen. To further enhance expression of the CTLA-4 VLD anti-lysozyme, the coding sequence was adjusted by splice overlap PCR to comprise codons preferential for *E.coli* expression.

Example 5

CTLA-4 STMs Based Upon a Human anti-Melanoma Antibody.

The human-derived anti-melanoma antibody V86 specifically binds human melanoma cells. This antibody is unusual in that binding affinity resides entirely within the V_H region, addition of a cognate V_L decreases binding efficiency, and that the V_H domain expressed with a small segment

of the V_L domain has a high degree of solubility (Cai and Garen, 1997). To further illustrate that replacement of CTLA-4 VLD CDR loop structures enhances solubility and that the resultant STMs can be produced in bacterial expression systems, the three CDR loop structures of CTLA-4 were replaced with the three CDR loop regions from V86. Positions and sequence of the substitutions are shown in Figure 6. Expression of this STM (3E4) in pGC again resulted in production of predominantly monomeric soluble protein (Figure 7) with enhanced solubility compared to the CTLA-4 VLD.

Example 6

Construction of CTLA-4 STMs as Libraries of Binding Molecules

To select CTLA-4 STMs with novel binding specificities, VLD libraries were produced containing randomised CDR1 and CDR3 loop structures. Oligonucleotide primers used for library construction are listed in figure 1. Combinations of oligonucleotide primers used for library construction are shown in Table 3.

Table 3. CTLA-4 STM Library Combinations

CDR1

CDR3	4483*	4254	5449	5451	5452	5450	5446	4835
4482	+1	+1	//////////	//////////	//////////	//////////	//////////	//////////
4275	+1	+1	//////////	//////////	//////////	//////////	//////////	//////////
5470	//////////	//////////	+2	+2	+2	+2	+2	//////////
5474	//////////	//////////	+2	+2	+2	+2	+2	//////////
5471	//////////	//////////	+2	+2	+2	+2	+2	//////////
5472	//////////	//////////	+2	+2	+2	+2	+2	//////////
5475	//////////	//////////	+2	+2	+2	+2	+2	//////////
5473	//////////	//////////	+2	+2	+2	+2	+2	//////////
4836	//////////	//////////	//////////	//////////	//////////	//////////	//////////	+3

*: oligonucleotide number.

+ : library combination.

1,2,3: describes library number.

DNA constructs encoding the resultant libraries were cloned into vectors pHFA or pFAB.5c for production of fd-phagemid based libraries and into pfd-Tet-Dog for production of fd-phage based libraries (see examples 1 and 2). Library 1 was cloned into vector pHFA and consisted of 2.1×10^7 independent clones. Library 3 was cloned into vectors pHFA (5.7×10^5 independent clones) and pfd-Tet-Dog (2.2×10^4 independent clones). Library 2 was cloned into pFAB.5c (1.7×10^7 independent clones) and into pfd-Tet-Dog (1.6×10^5 independent clones). Numbers of independent clones were determined by counting gross numbers of transformed colonies constituting the library, followed by assaying for the presence and proportion of CTLA-4 STM-specific DNA.

For library 2, the variability of the full library was tested by sequencing of representative clones. These results are presented in Table 4. The expected heterogeneity of insert size and sequence was observed. A high proportion of UAG termination codons were observed, consistent with the oligonucleotide randomisation strategy. To prevent these codons causing premature termination of the CTLA-4 STM gene3 protein fusions, libraries were transferred into the *E.coli* strains Tg-1 and JM109, which suppress this termination codon by insertion of a glutamic acid residue. Cysteine residues were present in the high numbers expected from the desgn of the oligonucleotides,and were in positions capable of forming intra- and inter-CDR loop structure disulphide bonds.

Table 4 CDR1 and CDR3 Inserts from a Representative Series of Library 2 Clones (SEQ ID NOS 102-138)

CLONE	CDR1	CDR3
3M-2	ND1	LPSSDTRAYS (SEQ ID NO: 102)
3M-3	QESGGRPG SEQ ID NO: 103	LPRGPPLLSL SEQ ID NO: 104
3M-5	SPGRCLN SEQ ID NO: 105	ND
3M-6	EWKR*HGG SEQ ID NO: 106	LCPGACGCVY SEQ ID NO: 107
3M-7	NSG*NEGG SEQ ID NO: 108	ND
3M-11	DKPVTKSG SEQ ID NO: 109	ND
3M-17	SPGACP* SEQ ID NO: 110	ND
3M-18	SPGKCDQ SEQ ID NO: 111	ND
3M-19	SPGMCAR SEQ ID NO: 112	LMYPPPYYL SEQ ID NO: 63
3M-20	ND	PFLFLPC*FFF SEQ ID NO: 113
3N-1	WTLGHHKLCEG SEQ ID NO: 114	LFTCLLALCS SEQ ID NO: 115
3N-2	SPGECYG SEQ ID NO: 116	SWLSTTXCLSSCS SEQ ID NO: 117
3N-3	SPG*CQD SEQ ID NO: 118	LLGSLLSCFASLS SEQ ID NO: 119
3N-4	SPG*CQD	SPGSLLSCFASXS SEQ ID NO: 120
3N-5	SPGRCTD SEQ ID NO: 121	VICHSSVCLSD/EVC SEQ ID NOS: 122 and 123
3N-6	ND	DLPSYLACSI SEQ ID NO: 124
3N-7	SPGRCDA SEQ ID NO: 125	ALCWDVFYCSFPSY SEQ ID NO: 126
3N-8	ELFGHARYCKG SEQ ID NO: 127	VSITSP*LCPVKVFD SEQ ID NO: 128
3N-9	SPGKV*N SEQ ID NO: 129	LFVPFVSP SEQ ID NO: 130
3N-12	SPGDLWV SEQ ID NO: 131	ESGLSPVSPCSLYSL SEQ ID NO: 132
3N-13	TSANGPYG SEQ ID NO: 133	PWAYRFLAVL SEQ ID NO: 134
3N-14	RKTREKYG SEQ ID NO: 135	ELMYPPPYYLGI SEQ ID NO: 136
3N-15	SPGQELT SEQ ID NO: 137	ELFFLLYAPCYLFQR SEQ ID NO: 138

ND: Not Done

*: UAG termination codon

Bacteriophage particles displaying CTLA-4 STMs as gene 3 protein fusions were rescued from *E.coli* cells by standard protocols and panned against antigens presented in a number of contexts as described in the following examples.

Example 7

CTLA-4 STM Libraries: Selection against Antigens on Solid Supports.

Four different antigens falling into a class of proteins with clefts or crevices within their structures were selected for screening. It was anticipated that the CTLA-4 VLD STMs, being of smaller size than antibodies, and possessing elongated CDR loop structures (especially CDR-3) would be able to access these cleft regions. The antigens selected were: (i) hen egg lysozyme (EC 3.2.1.17); (ii) bovine carbonic anhydrase (EC 4.2.1.1); (iii) fungal a-amylase (EC 3.2.1.1); and (iv) *Streptoalloteichus hindustanis* resistance protein ShBle (Gatignol et al. 1988). For binding to plates, antigens in coating buffer (1mg/ml in 0.1M NaHCO3 pH8.5) were bound to Costar ELISA plates by standard procedures. Rescued phage and phagemid-derived libraries were panned by standard and well-understood procedures except that lower than standard number of washes were employed to allow low affinity binding phage to be selected. Figure 11 shows titres of libraries selected against ShBle. After round 4, recovered bacteriophage titres were higher than controls. To those skilled in the art, this represents selection of specific binding moieties, and it is then a routine process to produce these selected CTLA-4 VLD STMs using expression vectors such as pGC or pPOW (as described in example 2).

Example 8

CTLA-4 STM Libraries: Selection against Antigens in Solution.

For selection in solution, the antigens bovine carbonic anhydrase and fungal a-amylase were biotinylated and selections performed in solution using capture by streptavidin coated magnetic beads. Throughout these experiments washes were kept constant at either 2 or 5 washes per selection round. Titres of recovered bacteriophage post-elution are shown in Figure 12. After round 4, recovered bacteriophage titres were higher than controls. To those skilled in the art, this represents selection of specific binding

moieties, and it is then a routine process to produce these selected CTLA-4 VLD STMs using expression vectors such as pGC or pPOW (as described in example 2).

Example 9

CTLA-4 STM Libraries: Selection in an Alternative Display and Selection System.

To allow further maturation and selection of antigen binding STMs, the CTLA-4 STM library was ligated into a plasmid to add a downstream C-terminal spacer polypeptide (heavy constant domain). Upstream transcriptional and translational initiation sequences were added by PCR amplification using specific oligonucleotides (Figure 1). This PCR DNA was used as a template for production of RNA followed by translation and display of the library on ribosomes in a coupled cell free translation system as described by He and Taussig (1997). To demonstrate binding, CTLA-4 STM ribosome complexes were panned on hepatitis B surface antigen (hbsa), glycophorin (glyA) and bovine serum albumin (BSA) coated dynabeads. RNA from ribosome complexes bound to hbsa, glyA and BSA was recovered by RT-PCR. It is then a routine process to clone these RT-PCR products into an expression vector such as pGC or pPOW (as described in example 2) allowing production of CTLA-4 STMs. It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention allowing display of libraries of CTLA-4 STMs as ribosome complexes (as in this example) as well as display on the surface of live cells (which may be derived from a eukaryotic or prokaryotic background) and may include bacteria, yeast, mammalian or insect cells.

Example 10

CTLA-4 STMs: Affinity Maturation and CDR2 Mutation.

To allow further maturation and selection of antigen-binding STMs, and the construction of randomised CDR-1, -2 and –3 libraries, CDR-2 randomised oligonucleotide primers were produced (Figure 1). A variation of these primers contained conserved cysteine residues to allow construction of STMs with CDR2-CDR3 disulphide linkages mimicing those found in llama single domain antibodies. Splice overlap PCR allowed creation of libraries containing all three CDR loop structures randomised.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

REFERENCES

Bodian DL, Jones EY, Harlos K, Stuart DI, Davis SJ (1994) Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 A resolution. Structure 2: 755-766

Cai X and Garen A (1997) Comparison of fusion phage libraries displaying VH or single-chain Fv antibody fragments derived from the antibody repertoire of a vaccinated melanoma patient as a source of melanoma-specific targeting molecules Immunol 94: 9261-9266

Canfield VA, Norbeck L, Levenson R (1996) Localization of cytoplasmic and extracellular domains of Na,K-ATPase by epitope tag insertion Biochem. 35: 14165-14172

Coia, G., Hudson, P.J., Lilley, G.G. (1996) Construction of recombinant extended single-chain antibody peptide conjugates for use in the diagnosis of HIV-1 and HIV-2. J. Immunol. Meth. 192:13-23

Davies J, Riechmann L (1994) Camelising human antibody fragments: NMR studies on V_H domains FEBS Lett. 339(3):285-90

Davies J, Riechmann L. (1996) Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human V_H domains with improved protein stability Protein Eng 9(6):531-7

Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S, Wyns L (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme Nat Struct Biol 3:803-811

Galanis M, Irving RA, and Hudson PJ (1997) Current Protocols in Immunology, 17.1.1-17.1.45.

Gatignol A, Durand H, and Tiraby G (1988) Bleomycin resistance conferred by a drug-binding protein FEBS Lett 230: 171-175

Gerstmayer B, Pessara U, Wels W (1997) Construction and expression in the yeast Pichia pastoris of functionally active soluble forms of the human costimulatory molecules B7-1 and B7-2 and the B7 counter-receptor CTLA-4. FEBS Lett 407: 63-68

Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R (1993) Naturally occurring antibodies devoid of light chains. Nature 363: 446-448

Hanes J and Pluckthun A (1997) In vitro selection and evolution of functional proteins by using ribosome display Proc.Natl.Acad. Sci. USA. 94: 4937-4942

He M and Taussig MJ (1997) Antibody-ribosome-mRNA (ARM) complexes as efficient selection particles for in vitro display and evolution of antibody combining sites. Nucl. Acids Res. 25: 5132-5134

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., and Winter, G. 1991. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. 19:4133-4137.

Huse, W. D.,Sastry, L.,Iverson, S. A.,Kang, A. S.,Alting, M. M.,Burton, D. R.,Benkovic, S. J., & Lerner, R. A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda . Science, 246: 1275-81.

Hutloff, A., Dittrich, A.M., Beier, K.C., Elijaschewitsch, B., Kraft, R., Anagnostopoulos, I. and Kroczek, R.A. (1999). ICOS is an inducible T-cell co-stimulator structurally and functionally related to CD28. Nature 397: 263-266.

Kortt, A.A., Guthrie, R.E., Hinds, M.G., Power, B.E., Ivancic, N., Caldwell, J.B., Gruen, L.C., Norton, R.S. and Hudson, P.J. (1995) Solution properties of E. coli expressed VH domain of anti-neuraminidase antibody NC41. J. Protein Chemistry. 14, 167-178.

Linsley, PS, Clark EA, and Ledbetter JA (1990) The T cell antigen CD28 mediates adhesion with B cells by interacting with the activation antigen, B7/BB1 Proc. Natl. Acad. Sci USA 87: 5031.

Linsley PS, Nadler SG, Bajorath J, Peach R, Leung HT, Rogers J, Bradshaw J, Stebbins M, Leytze G, Brady W, et al (1995) Binding stoichiometry of the cytotoxic T lymphocyte-associated molecule-4 (CTLA-4). A disulfide-linked homodimer binds two CD86 molecules. J Biol Chem 270: 15417-24

McCafferty, J..Griffiths, A. D.,Winter, G., & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. Nature, 348: 552-4.

Metzler WJ, Bajorath J, Fenderson W, Shaw SY, Constantine KL, Naemura J, Leytze G, Peach RJ, Lavoie TB, Mueller L, Linsley PS (1997) Solution structure of human CTLA-4 and delineation of a CD80/CD86 binding site conserved in CD28 Nat Struct Biol 4: 527-531

Morton PA, Fu XT, Stewart JA, Giacoletto KS, White SL, Leysath CE, Evans RJ, Shieh JJ, Karr RW (1996) Differential effects of CTLA-4 substitutions on the binding of human CD80 (B7-1) and CD86 (B7-2) J Immunol 156: 1047-1054

Muyldermans S, Atarhouch T, Saldanha J, Barbosa JA, Hamers R (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Eng 7: 1129-1135.

Nieba L, Honegger A, Krebber C, Pluckthun A (1997) Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment Protein Engineering (4):435-44

Novotny J, Ganju RK, Smiley ST, Hussey RE, Luther MA, Recny MA, Siliciano RF, Reinherz EL (1991) A soluble, single-chain T-cell receptor fragment endowed with antigen-combining properties Proc Natl Acad Sci U S A 88(19):8646-8650.

Peach RJ, Bajorath J, Brady W, Leytze G, Greene J, Naemura J, Linsley PS (1994) Complementarity determining region 1 (CDR1)- and CDR3-analogous regions in CTLA-4 and CD28 determine the binding to B7-1 J Exp Med 180: 2049-2058

Power, B.E., Ivancic, N., Harley, V.R., Webster, R.G., Kortt, A.A, Irving, R.A. and Hudson, P.J. (1992) High-level temperature-induced synthesis of an antibody VH-domain in *Escherichia coli* using the PelB secretion signal gene 113:95-99.

Reisine, T. (1995) Somatostatin. Cell Molec. Neurobiol. 15: 597-614

Reubi, J. C. (1997) Regulatory peptice receptors as molecular targets for cancer diagnosis and therapy. Q. J. Nucl. Med. 41: 63-70

Schonbrunn A, Gu YZ, Brown PJ, Loose-Mitchell D (1995) Function and regulation of somatostatin receptor subtypes. Ciba Found. Symp. 190: 204-217

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science, 228, 1315-1317

van der Merwe PA, Bodian DL, Daenke S, Linsley P, Davis SJ (1997) CD80 (B7-1) binds both CD28 and CTLA-4 with a low affinity and very fast kinetics. J Exp Med 185: 393-403

Vu KB, Ghahroudi MA, Wyns L, Muyldermans S (1997) Comparison of llama VH sequences from conventional and Heavy Chain Antibodies. Molec Immunol 34: 1121-1131

Ward ES (1991) Expression and secretion of T-cell receptor V alpha and V beta domains using *Escherichia coli* as a host Scand J Immunol 34(2):215-220.

Waterhouse P, Marengere LE, Mittrücker HW, Mak TW (1996) CTLA-4, a negative regulator of T-lymphocyte activation. Immunol Rev 153: 183-207

Waterhouse P, Penninger JM, Timms E, Wakeham A, Shahinian A, Lee KP, Thompson CB, Griesser H, Mak TW (1995) Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4 Science 270: 985-988

Winter G, Milstein C. (1991). Man-made antibodies. Nature. 349: 293-299.

Wulfing C, Pluckthun A and in (1994) Correctly folded T-cell receptor fragments in the periplasm of *Escherichia coli* - Influence of folding catalysts J Mol Biol 242(5):655-69.